UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/770,092 | 02/02/2004 | Tae-jin Ahn | YPL-0078 | 9277 |

23413 7590 01/09/2008

CANTOR COLBURN, LLP
20 Church Street
22nd Floor
Hartford, CT 06103

| EXAMINER | |
|---|---|
| SMITH, CAROLYN L | |
| **ART UNIT** | **PAPER NUMBER** |
| 1631 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 01/09/2008 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/770,092 | AHN, TAE-JIN | |
| | Examiner | Art Unit | |
| | Carolyn L. Smith | 1631 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *15 October 2007 and 17 September 2007*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-3 and 5-19* is/are pending in the application.

4a) Of the above claim(s) *10-18* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-3, 5-9, 19* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

## DETAILED ACTION

A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicants' submission, filed 10/15/07 and 9/17/07, have been entered.

Amended claims 1-3, 9, and 19 and cancelled claim 4, filed 9/17/07, are acknowledged. Claims 10-18 remain withdrawn due to being drawn to a non-elected Group.

Claims herein under examination are 1-3, 5-9, and 19.

### *Claim Rejections – 35 USC § 103*

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-3, 5-7 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Grumbach et al. (Information Processing & Management, Volume 30, Number 6, pages 875-886, 1994) in view of Robson et al. (1992).

Grumbach et al. describe using UNIX (computer operating system) and ASCII files with algorithms to encode DNA sequences (page 876, second and third paragraphs and page 878, second and third paragraphs) as well as data compression for storage of sequences (page 875, first paragraph) and compression rates (page 878, fourth paragraph to page 879, second paragraph) which represents an apparatus and computer readable medium for a encoding DNA sequence to achieve a high data compression ratio for storage, as stated in the preamble of instant claims 1 and 19. Grumbach et al. describe storing a DNA sequence by encoding individual DNA characters (Figure 1 and page 878, third paragraph). Grumbach et al. describe using a dictionary containing already encoded factors (page 879, third paragraph) that represents predetermined conversion codes. Grumbach et al. describe having a start position and matching the factor at the current position followed by outputting a codeword (page 881, second paragraph), alignment algorithms working directly on compressed DNA sequences (page 886, first paragraph), compression algorithms including a vertical mode, where a DNA sequence A is compressed with respect to another sequence B with output containing information to construct sequence A from sequence B (page 876, fifth paragraph), using codewords to encode strings and arithmetic encoding (page 876, third paragraph and page 879, first paragraph), and storing a reference sequence in a database while other sequences are stored in a compressed form with respect to it (page 876, fifth paragraph)as well as alignment algorithms working directly on compressed DNA sequences (page 886, first paragraph) which represents the comparative, conversion, code storage, and encoding units (as stated in instant claim 1), the conversion (as stated in instant claim 6), the compression and storage units (as stated in instant claim 7), and the aligning, extracting, converting, and encoding steps (as stated in instant claim 19). Grumbach et al.

describe statistical and substitutional compression of text via encoding blocks of fixed length and encoding factors of different lengths using a pointer to one of their previous occurrences in the text (page 875, third paragraph) which represents a division unit, as stated in instant claim 6. Grumbach et al. describe using a codeword including "l" for the length of the factor, and "p" for the position of the first occurrence (page 881, first paragraph). Grumbach et al. describe encoding DNA base symbols on bits of codewords (page 881, last paragraph) and characters representing the number of substitutional difference during compression (page 882, first and second paragraphs). Grumbach et al. describe using output code containing three types of codewords: literal, numerical, and copy codewords (page 882, first paragraph). Grumbach et al. describe using characters to code DNA symbols (page 881, last paragraph) and characters representing the number of substitutional difference during compression (page 882, first and second paragraphs). Grumbach et al. describe outputting codewords (page 881, second paragraph). Grumbach et al. do not describe all of the type of extracted difference, characters representing the extracted difference, extracted difference features, or 4-bit codes as stated in instant claims 2-3, 5, and 19.

Robson et al. describe using computers and computer programs (page 285, column 1, second paragraph) which inherently contain RAM and ROM which represents a computer readable medium. Robson et al. describe storing and analyzing nucleic acid sequences including quantifying and exploring sequence relationships, using standard searching algorithms (i.e. BLAST), comparing sequences (page 285, column 1, second paragraph and page 288, first paragraph), encoding sequence information (page 283, first column, last paragraph and second column, third and fourth paragraphs), assigning code based on a sub-selection (i.e. extracted

difference) defined by classes of properties (abstract), as well as sorting, searching, and abstracting sub-sets (page 283, col. 2, last paragraph). Robson et al. describe using 4 bit code (page 286, first column, line 2; page 287, first column, last paragraph), as stated in instant claim 5. Robson et al. describe using 4 bit codes which correspond to characters including the start or end of the sequence, continue to read, ending, sequence separator (i.e. distance between start and end position), matches, codes for purines and pyrimidines (i.e.A, G, C), and various other characters (page 287, column 1, last paragraph to column 2, first paragraph), as stated in instant claims 2 and 3. Robson et al. describe using code to signify differences and their number, such as unknowns, blanks, and deletions (page 287, first column), as stated in instant claims 1-3 and 19.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the apparatus of Grumbach et al. by using the 4 bit code words and characters as taught by Robson et al. where the motivation would have been to perform searching in a more intelligent, structured, and faster manner since numerical bioinformatics descriptions are of value whenever the quality and quantity of information is very large, as stated by Robson et al. (page 284, column 1, third paragraph and abstract).

Thus, Grumbach et al., in view of Robson et al., make obvious claims 1-3, 5-7, and 19.

Applicant summarizes the rejection and argues that Grumbach et al. do not teach an algorithm wherein the extracted difference comprises a start region mismatch, blank, single base pair mismatch, base insertion, multiple pair mismatch, or an end region mismatch. This statement is found unpersuasive as Robson et al. teach this limitation. Applicant summarizes

Robson et al. and argues that Robson et al. do not teach converting an extracted difference between the reference sequence and the subject sequence into a string of characters. This statement is found unpersuasive as Grumbach et al. describe using a dictionary containing already encoded factors (page 879, third paragraph) that represents predetermined conversion codes. Grumbach et al. disclose having a start position and matching the factor at the current position followed by outputting a codeword (page 881, second paragraph), alignment algorithms working directly on compressed DNA sequences (page 886, first paragraph), compression algorithms including a vertical mode, where a DNA sequence A is compressed with respect to another sequence B with output containing information to construct sequence A from sequence B (page 876, fifth paragraph), and using codewords to encode strings and arithmetic encoding (page 876, third paragraph and page 879, first paragraph). Applicant argues that Robson et al. do not describe that the extracted difference comprises a start region mismatch between the reference sequence and the subject sequence; a blank representing there is no base in a base position in the subject sequence corresponding to the reference sequence; a single base pair mismatch between the reference sequence and the subject sequence; a base insertion into the subject sequence; a multiple base pair mismatch between the reference sequence and the subject sequence, or an end region mismatch between the reference sequence and the subject reference. This statement is found unpersuasive as Robson et al. describe using 4 bit codes which correspond to characters including the start or end of the sequence, continue to read, ending, sequence separator (i.e. distance between start and end position), matches, codes for purines and pyrimidines (i.e.A, G, C), and various other characters (page 287, column 1, last paragraph to column 2, first paragraph). Robson et al. describe using code to signify differences and their

number, such as unknowns, blanks, and deletions (page 287, first column). Applicant's arguments are deemed unpersuasive for the reasons given above.

Claims 8-9 are rejected under 35 U.S.C. 103(a) as being unpatentable over Grumbach et al. (Information Processing & Management, Volume 30, Number 6, pages 875-886, 1994) in view of Robson et al. (1992) as applied to claims 1-3, 5-7, and 19 above, and further in view of Selifonov et al. (US 2002/0183934 A1).

Grumbach et al. and Robson et al. describe the limitations of instant claims 1-3, 5-7 and 19, as discussed in the 35 USC 103 rejection above. Grumbach et al. describe using characters to code DNA symbols (page 881, last paragraph) and characters representing the number of substitutional difference during compression (page 882, first and second paragraphs). Grumbach et al. describe outputting codewords (page 881, second paragraph). Grumbach et al. and Robson et al. do not describe all of the limitations stated in instant claims 8 and 9.

Selifonov et al. describe making character strings for polynucleotides (title). Selifonov et al. describe modifying a parental character string sequence (0025, claim 17) and generating random variation of sequences via multiplication factors (0075-0076, 0117, 0197), as stated in instant claim 8. Selifonov et al. describe sequences including lengths, type, number/round of evolution and nucleic acid shuffling, and mutated fragments of predefined lengths, and software for sequence string manipulation (0026-0028) as well as mutation types from a parent sequence, mutation (i.e. single point mutation [total number of variation = 1], continuous mutation over entire string, triple deletion frameshifts), fragmentation, crossover, ligation, and elitism with

string analysis tools (i.e. sequence length, specific substrings) (0075-0080) which represents the variation sequence generation factor, as stated in instant claim 9. Selifonov et al. describe using a computer readable medium and a CD-ROM (0133 and 0286).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the apparatus of Grumbach et al. by using the 4 bit code words and characters as taught by Robson et al. where the motivation would have been to perform searching in a more intelligent, structured, and faster manner since numerical bioinformatics descriptions are of value whenever the quality and quantity of information is very large, as stated by Robson et al. (page 284, column 1, third paragraph and abstract). It would have been further obvious to one of ordinary skill in the art at the time the invention was made to modify the apparatus of Grumbach et al. and Robson et al. by creating a variation sequence as taught by Selifonov et al. wherein the motivation would have been to provide for the rapid evolution of nucleic acids for the generation of encoded molecules (e.g., nucleic acids and proteins) with new and/or improved properties of industrial, agricultural and therapeutic importance which can be created or improved through DNA shuffling procedures, as stated by Selifonov et al. (0008).

Thus, Grumbach et al. in view of Robson et al. and Selifonov et al. make obvious claims 1-3, 5-9, and 19.

Applicant summarizes Selifonov et al. Applicant argues that Selifonov et al. do not teach an algorithm to extract the difference between a subject sequence and a reference sequence wherein the extracted difference comprises a start region mismatch between the reference sequence and the subject sequence; a blank representing there is no base in a base position in the

subject sequence corresponding to the reference sequence; a single base pair mismatch between the reference sequence and the subject sequence; a base insertion into the subject sequence; a multiple base pair mismatch between the reference sequence and the subject sequence, or an end region mismatch between the reference sequence and the subject sequence. This statement is found unpersuasive as Grumbach et al. and Robson et al. describe these limitations (see above), and Applicant has failed to provide a solid reasoning as to why the passages in Grumbach et al. and Robson et al. would be considered improper. Applicant's arguments are deemed unpersuasive for the reasons given above.

***Conclusion***

No claim is allowed.

Papers related to this application may be submitted to Technical Center 1600 by facsimile transmission. Papers should be faxed to Technical Center 1600 via the PTO Fax Center. The faxing of such papers must conform to the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993) (See 37 CFR §1.6(d)). The Central Fax Center number for official correspondence is (571) 273-8300.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Carolyn Smith, whose telephone number is (571) 272-0721. The examiner can normally be reached Monday through Thursday from 8 A.M. to 6:30 P.M.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marjorie Moran, can be reached on (571) 272-0720.

January 3, 2008

/Carolyn Smith/
Primary Examiner
AU 1631